Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **510,** 2851–2863 (2022)
Advance Access publication 2021 December 9

https://doi.org/10.1093/mnras/stab3407

A *Swift* study of long-term changes in the X-ray flaring properties of Sagittarius A*

A. Andrés,[1,2,3★] J. van den Eijnden,[1,4★] N. Degenaar,[1] P. A. Evans,[5] K. Chatterjee,[1,6,7] M. Reynolds,[8] J. M. Miller,[8] J. Kennea,[9] R. Wijnands,[1] S. Markoff,[1] D. Altamirano,[10] C. O. Heinke,[11] A. Bahramian,[12] G. Ponti[13,14] and D. Haggard[15]

[1]*Anton Pannekoek Institute for Astronomy, University of Amsterdam, Science Park 904, NL-1098 XH Amsterdam, the Netherlands*
[2]*Instituto de Astronomía, Universidad Nacional Autonoma de México, Av. Universidad 3000, C.U., Coyoacán, 04510 Ciudad de México, CDMX, México*
[3]*Facultad de Ciencias Naturales y Matemática, Universidad de El Salvador, Final 25 Avenida Norte, 01101 San Salvador, El Salvador*
[4]*Department of Physics, Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[5]*School of Physics and Astronomy, University of Leicester, University Road, Leicester LE1 7RH, UK*
[6]*Black Hole Initiative at Harvard University, 20 Garden Street, Cambridge, MA 02138, USA*
[7]*Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[8]*Department of Astronomy, University of Michigan, 1085 South University Avenue, Ann Arbor, MI 48109, USA*
[9]*Department of Astronomy and Astrophysics, Pennsylvania State University, 525 Davey Lab, University Park, PA 16802, USA*
[10]*Department of Physics and Astronomy, University of Southampton, Southampton, Hampshire SO171BJ, UK*
[11]*Department of Physics, University of Alberta, 4-183 CCIS, Edmonton, AB T6G 2E1, Canada*
[12]*International Centre for Radio Astronomy Research, Curtin University, GPO Box U1987, Perth, WA 6845, Australia*
[13]*Osservatorio Astronomico di Brera, Via E. Bianchi 46, I-23807 Merate (LC), Italy*
[14]*Max-Planck-Institut für Extraterrestrische Physik, Giessenbachstrasse, D-85748 Garching, Germany*
[15]*Department of Physics, McGill University, 3600 rue University, Montreal, QC H3A 2T8, Canada*

Accepted 2021 November 16. Received 2021 November 9; in original form 2021 July 27

ABSTRACT

The radiative counterpart of the supermassive black hole at the Galactic Centre, Sagittarius A*, displays flaring emission in the X-ray band atop a steady, quiescent level. Flares are also observed in the near-infrared band. The physical process producing the flares is not fully understood and it is unclear if the flaring rate varies, although some recent works suggest it has reached unprecedented variability in recent years. Using over a decade of regular X-ray monitoring of Neil Gehrels *Swift* Observatory, we studied the variations in count rate of Sgr A* on time-scales of years. We decomposed the X-ray emission into quiescent and flaring emission, modelled as a constant and power-law process, respectively. We found that the complete, multiyear data set cannot be described by a stationary distribution of flare fluxes, while individual years follow this model better. In three of the ten studied years, the data is consistent with a purely Poissonian quiescent distribution, while for 5 yr, only an upper limit of the flare flux distribution parameter could be determined. We find that these possible changes cannot be explained fully by the different number of observations per year. Combined, these results are instead consistent with a changing flaring rate of Sgr A*, appearing more active between 2006–2007 and 2017–2019, than between 2008–2012. Finally, we discuss this result in the context of flare models and the passing of gaseous objects, and discuss the extra statistical steps taken, for instance, to deal with the background in the *Swift* observations.

Key words: black hole physics – Galaxy: centre – X-rays: individual: (Sgr A*).

1 INTRODUCTION

Sagittarius A* (Sgr A*) is the electromagnetic counterpart of the supermassive black hole at the centre of the Milky Way galaxy. It has an estimated mass of $\sim 4 \times 10^6\,M_\odot$, but its bolometric luminosity is ∼9 orders of magnitude fainter than the Eddington luminosity for an object of this mass (Genzel, Eisenhauer & Gillessen 2010; Morris, Meyer & Ghez 2012). It is the most nearby galactic nucleus, located at a distance of ∼8 kpc from Earth (Reid & Brunthaler 2004; Ghez et al. 2008), this makes Sgr A* the prime laboratory to study the accretion processes on to supermassive black holes at such low accretion rates.

The X-ray emission from Sgr A* is observed to be composed of a quiescent component, corresponding to a luminosity of $L_X \simeq 3 \times 10^{33}\,\mathrm{erg\,s^{-1}}$ in the 2–10 keV energy range, which is interrupted almost daily by flares (e.g. Baganoff et al. 2001; Goldwurm et al. 2003; Genzel et al. 2010; Markoff 2010; Degenaar et al. 2013; Neilsen et al. 2013). These flares are ∼1−2 orders of magnitude more luminous than its quiescent emission, with the brightest ones reaching values of $L_X \simeq (1-5) \times 10^{35}\,\mathrm{erg\,s^{-1}}$ (e.g. Nowak et al. 2012;

★ E-mail: bandres@astro.unam.mx (AA); jakob.vandeneijnden@physics.ox.ac.uk (JVDE)

Haggard et al. 2019). Sgr A* is also flaring in other wavebands, most prominently in the near-infrared (nIR; e.g. Witzel et al. 2018). Both, the emission mechanism and the physical process producing Sgr A*'s flares are not completely understood yet (Markoff et al. 2001; Liu & Melia 2002; Yuan, Quataert & Narayan 2003; Liu, Petrosian & Melia 2004; Čadež, Calvani & Kostić 2008), although Ponti et al. (2017) developed the first simultaneous multiwaveband campaign measuring the spectral index in nIR and X-ray bands, showing that synchrotron emission with a cooling break is a viable process for Sgr A*'s flaring emission.

In the past decade, extensive work has been performed to simultaneously detect flares from Sgr A* in different wavebands in order to gain more insight into the emission mechanism (Eckart et al. 2004; Yusef-Zadeh et al. 2008; Trap et al. 2011). Another avenue of study has been to characterize the brightness distribution and occurrence rate of flares at X-ray and nIR wavelengths. In the X-rays, the first systematical work was developed before the passage of the G2 object to the Galactic Centre (for reviews of the G2 object, see Gillessen et al. 2012; Witzel et al. 2014). Using 3Ms of data from the *Chandra X-ray Observatory's* 2012 X-ray Visionary Project (XVP), Neilsen et al. (2013) reported 39 flares for this set of data and a flaring rate of ∼1.1 flares per day, which is consistent with previous results by Genzel et al. (2010). It also has been shown that Sgr A* X-ray flux distribution can be decomposed into the sum of two processes: a quiescent component with constant flux, and a flaring component best described by a power-law distribution of fluxes (Neilsen et al. 2015). Regarding the nIR band, several works have also been performed. For instance, by making use of the GRAVITY instrument, Abuter et al. (2018) detected orbital motions of three NIR flares from Sgr A*. Also, using data from the *Keck* Observatory, *Spitzer* Space Telescope, and the *Very Large Telescope (VLT)*, Witzel et al. (2018) found that the variability of Sgr A* in this band can be described as a red noise process with a single lognormal distribution.

Establishing if the flaring properties of Sgr A* change over time could provide new hints into the physical mechanism producing the flares. For instance, Ponti et al. (2015) claimed evidence for a change in Sgr A* flaring rate, using 6.9 Ms of *Chandra* and *XMM–Newton* data, covering 1999 September–2014 November. They reported a significant increase in the rate of the very bright flares in late 2013 and 2014, changing from 0.27 ± 0.04 to 2.5 ± 1.0 flares per day at a 99.9-per cent confidence level. On the other hand, Bouffard et al. (2019) studied a total of 4.5 Ms of *Chandra* observations only, covering 2012–2018 data free of contamination from the magnetar SGR J1745–2900 (see Section 2.1), but did not find evidence of change in Sgr A* flaring properties between the XVP and post-XVP data. Similarly, based on 4.5Ms of *Chandra* observations from 1999 to 2012, Yuan & Wang (2015) and Yuan et al. (2017) did not find evidence of changes in the quiescent nor flaring rate of Sgr A*, even around the pericentre passage of the S2 star in 2002. Another recent work suggests the emission of Sgr A* in the nIR band has been consistent during ∼20 yr of observations (Chen et al. 2019), but apparently, the emission has reached unprecedented flux levels in 2019, with flux peaks that are twice the values from previous measurements (Do et al. 2019).

In this work, we study the Cumulative Distribution Function (CDF) of count rates of Sgr A* using data accumulated between 2006 and 2019 with the Neil Gehrels *Swift* observatory (*Swift*; Gehrels et al. 2004). The long-term monitoring and high observing cadence of the *Swift* programme uniquely allow us to test whether the properties of Sgr A*'s X-ray flaring behaviour show evidence of changes on a time-scale of years.

2 OBSERVATIONS AND METHODS

2.1 *Swift*/XRT long-term light curve

In 2006 February, *Swift* started to monitor the Galactic Centre (GC) with the on-board X-Ray Telescope (XRT) with the aim of studying Sgr A* as well as numerous transient X-ray binaries located in this region (Kennea et al. 2006). Apart from a handful of interruptions and Sun constraints, *Swift*/XRT has pointed at the GC every ∼1–3 d since 2006, with an average exposure time of 1 ks per observation (see Degenaar et al. 2015 for a review of the program).

In this work, we used all available *Swift*/XRT data that covered Sgr A* and was obtained in photon-counting (PC) mode. The data spans the period between 2006 February 24 and 2019 August 6. The light curve of Sgr A* was extracted with the software implemented in the online XRT data analysis tool (Evans et al. 2007, 2009),[1] with the only exception that a fixed source and background region were used. To extract events from Sgr A*, we employed a circular region with a 10-arcsec radius centred at RA = 266.41682 and Dec. = −29.007797 (J2000). To account for the background, we used three circular regions of 10 arcsec that were free of X-ray point sources but did contain diffuse X-ray emission (as seen at the position of Sgr A*).

The long-term light curve was created with a bin size of 500 s and it is shown in Fig. 1. It clearly shows two bright, extended periods of activity that do not belong to Sgr A*, but to transient sources located within 10 arcsec of the supermassive black hole. The first is the transient magnetar SGR J1745–29, which exhibited an outburst between 2013 and 2015 (Kennea et al. 2013; Coti Zelati et al. 2017), and the second is the transient X-ray binary *Swift* J174540.7–290015 that was active in 2016 (Ponti et al. 2016; Reynolds et al. 2016). Given the brightness of these objects and the small angular separation of these from Sgr A*, compared to the point spread function of *Swift*, it is impossible to extract reliable information on the brightness of Sgr A* during the time that these transients were active. Therefore, we excluded all data obtained between 2013 March 31 and 2016 July 28 from our analysis of Sgr A*.

Apart from the obvious outbursts of the two transients above, the *Swift* light curve of Sgr A* shows some instances of elevated emission that have been ascribed to flares from Sgr A*. Based on the significance of these high points compared to the long-term average XRT count rate at the position of Sgr A*, several bright X-ray flares have been reported previously (Degenaar et al. 2013, 2015, 2019; Reynolds et al. 2018).

2.2 Modelling the count rate distribution

In this work, we focus on the count rate distribution, instead of individual flares. We followed a similar approach to the one described in Neilsen et al. (2015) to analyse the CDF of Sgr A*'s light curve, shown in the bottom panel of Fig. 3. The empirical CDF is the fraction of rates greater than or equal to a given rate, and is defined by the equation

$$f(r) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\{r_i \geq r\}}, \quad (1)$$

where N is the number of time bins, $\mathbb{1}$ is the indicator function[2], and r_i are the count rates. With this definition, $f(r)$ runs from one to

[1] https://www.swift.ac.uk/user_objects/

[2] The indicator function, $\mathbb{1}$, is defined to be 1 when the condition is true, and 0 elsewhere.

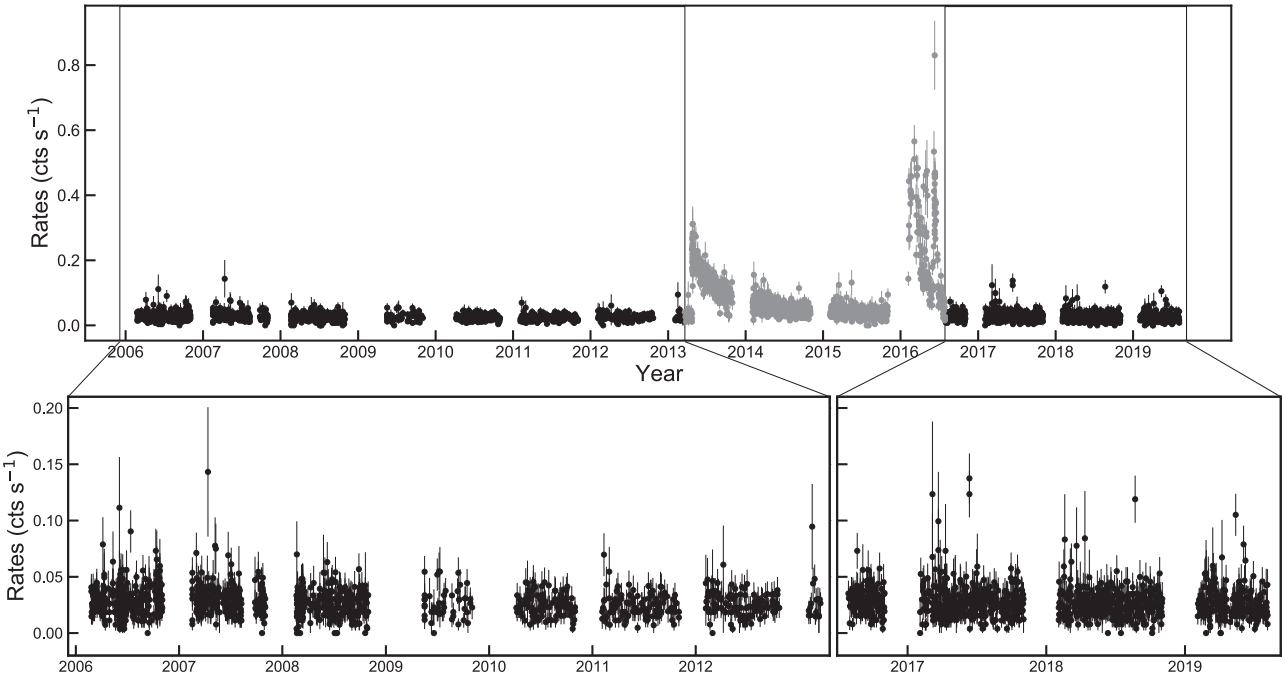

Figure 1. Long-term *Swift*/XRT-PC light curve of Sgr A* at 500-s binning (0.3–10 keV). The data marked in grey in top panel were excluded from our analysis due to the activity of nearby transient X-ray sources (see Section 2.1). The bottom panel shows the Sgr A* data that we used for the analysis.

zero, allowing for an easier visualization of changes at the high rate end on a logarithmic scale: $f(0) = 1$ per definition, as count rates are non-negative. $f(r_h) = 1/N$, where r_h is the highest observed count rate in our data, while $f(r) \equiv 0$ for any $r > r_h$.

Following the results from Neilsen et al. (2015), we assumed that the emission from Sgr A* is composed of a quiescent and a flaring emission component. We model these components as a constant and a power-law distribution in flux, respectively. We opted for the power-law distribution for the flux of the flaring component because Neilsen et al. (2015) found such a distribution provides the best description of 2012 *Chandra* XVP data – since the *Chandra* observations have a higher sensitivity then *Swift*, we do not expect to see deviations from this power-law model in the CDF of a similar-length light curve (i.e. ∼1 yr). To generate the flaring component, one can define a power-law probability distribution of fluxes straightforwardly as

$$P(F) = \begin{cases} kF^{-\alpha}, & F_{\min} < F < F_{\max} \\ 0, & \text{elsewhere}, \end{cases} \tag{2}$$

where F represents the flux, α the power-law index, and $k = (\alpha - 1)/(F_{\min}^{1-\alpha} - F_{\max}^{1-\alpha})$ is a constant of normalization.[3]

This definition, introduced by Neilsen et al. (2015), carries a significant downside, however, for data where flaring is either not present or faint in comparison to the quiescent rate and its uncertainties: the parameter characterizing the flaring rate, α, is unbounded. Low values of α, approaching zero, imply an average flare flux that is relatively high within the considered range $F_{\min}$ to $F_{\max}$. On the other hand, if *no* flaring is present, α will asymptotically tend to infinity as the flare flux tends to a constant value of $F_{\min}$ – combined, the quiescent and flaring components then form the equivalent of a single quiescent component. In reality, our analysis shows that this effect shows up already when $\alpha \approx 3$–4, which means that any $\alpha \geq 3$ will yield the same model fit and quality. As a result, for a light curve without (detectable) flaring, α will be unconstrained at high values.

To counteract this issue, we instead introduce a re-parametrization to our analysis, defined as

$$\zeta \equiv \frac{1}{\alpha}. \tag{3}$$

This new flaring parameter ζ is mathematically better constrained, with the lower bound 0 corresponding to the case of no flaring and its upper bound set by the distribution of flares in the data – arbitrarily high values of ζ are therefore not possible, as those correspond to ever-increasing flare fluxes. Moreover, ζ is more intuitively interpreted, with low values corresponding to low flare fluxes. We stress that this approach does not fundamentally differ from e.g. Neilsen et al. (2015); instead, it simply allows us to calculate a proper upper limit on the flaring parameter when no flares are observed, which is not possible with the unbounded α-parameter.[4]

Furthermore, we note that the simulated fluxes do not depend strictly on k, but only on $F_{\min}$, $F_{\max}$, and ζ. We fixed the values (in units of 10^{-12} erg cm^{-2} s^{-1}) to be $F_{\min} = 0.1$ and $F_{\max} = 16.0$ for the minimum and maximum flux, respectively. This $F_{\min}$ is higher than that from Neilsen et al. (2015), to account for the lower sensitivity of *Swift* and corresponds to its limiting sensitivity in 500 s. For the maximum, we took the brightest flare measured with *Swift* as reported by Degenaar et al. (2013).

In the case that no significant background emission is present (as can be done for, e.g. the *Chandra* XVP data), the flaring flux can simply be converted into a count rate assuming a certain flux-to-counts conversion and Poisson statistics. However, given the relatively high background in the *Swift* Galactic Centre data,

[3]This definition of k corrects a small typo in the original model from Neilsen et al. (2015).

[4]Also, our *Swift* analysis shows that α does not follow a Gaussian distribution, that would be distorted through this conversion. As such, our analysis does not lose any advantages that would come from assuming Gaussian distributions, due to this change of parameters.

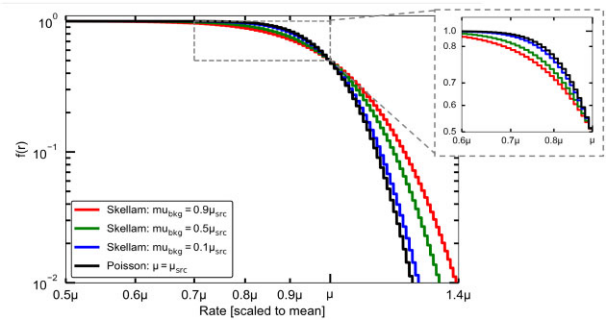

Figure 2. The CDF shape for one Poisson and three Skellam-distributed data sets with the same mean $\mu_{\rm src}$, assuming different relative background levels $\mu_{\rm bkg}$.

Poisson statistics do not hold. In the source region, we measure the sum of source counts and background counts, both of which are Poisson distributed. The former is then corrected by subtracting the background counts as measured in a separate region, after which the remaining source counts as modelled are the sum of quiescent and flaring emission. Assuming the background is spatially constant, this two-component model correctly describes the *mean* of the source counts, since the mean of two subtracted Poisson variables equals the subtraction of their respective means. However, the resulting *distribution* of source counts does not follow a Poisson but a Skellam distribution, defined by the total counts in the source region and those in the background region.

To highlight the difference between the Poisson and Skellam distributions, we plot four examples of the CDF $f(r)$ in Fig. 2. The three examples for Skellam distributions, with background rates at 90, 50, and 10 per cent of the mean count rate, show the significant distortion to the CDF shape. In the case of our *Swift* data, we find that the ratio $b/\mu \approx 0.1$, and hence we need to account for this distortion by explicitly modelling the data using the Skellam distribution. Otherwise, the CDF shape would be incorrectly reproduced and, possibly, the flaring flux would be overestimated.

Therefore, we model the distribution of source count rates, assuming a value of ζ and the quiescent count rate Q, using the equation

$$r(F, Q) = \frac{1}{\Delta t}\texttt{skellam}\left[\left(F \times \left(\frac{r_{\rm max}}{F_{\rm max}}\right) + Q + b\right)\Delta t, b\Delta t\right], \tag{4}$$

where $\Delta t = 500$ s is the bin size we used, `Skellam` is a Skellam random number generator, b is the measured background region count rate, F is the simulated flux taken from equation (2), and $r_{\rm max} = 10.2 \times 10^{-2}$ count s^{-1} is the mean observed count rate of the brightest flare (Degenaar et al. 2013). The Skellam random number generator takes the total source and background region counts as its two inputs, and returns a random draw of the former corrected by the latter. We applied a constant-rate conversion because no significant variations in the spectrum of Sgr A* flares have been observed (e.g. Degenaar et al. 2015). Furthermore, we are not applying a pile-up correction, because the count rates seen by *Swift*, even in flares, from Sgr A* are all below 0.2 count s^{-1}, i.e. well below the rates where this effect becomes important (see Fig. 1).

To fit the observed X-ray distribution of Sgr A*, we generated sets of synthetic data consisting of a combination of: a quiescent count rate distribution, with mean Q, and a power-law flux distribution of index $\zeta \equiv 1/\alpha$. To find the values of the parameters that best match the observations, we considered the two methods explained below.

2.2.1 Two-dimensional method

Following the same approach as Neilsen et al. (2015), in the two-dimensional (2D) method, we create a grid for values of ζ between 0 and 1, and values of Q between 14 and 26 count ks^{-1}, with 100 steps in both. Fits on larger grids show that outside these ranges, the probability of a match between the real data and synthetic data is null and/or unchanging.

For every pair (ζ, Q), we generated 1000 sets of synthetic distributions and applied the Anderson–Darling test (hereinafter called AD-test; Anderson & Darling 1954) to each combination of the real and a synthetic light curve to compare their distributions. At each parameter pair, we stored the average test-statistics (TS) from the AD-tests and then we converted it to p-values to assess the probability of both distributions to be drawn from the same underlying distribution. We then computed the median value and the 1σ confidence interval of ζ and Q from the marginalized probability contours generated in the $\zeta - Q$ plane; alternatively, if the distribution did not converge to zero at low ζ, we instead calculated the 90-per cent upper limit of ζ. We note that the measured values of ζ and Q depend on $F_{\rm min}$, since for larger values of $F_{\rm min}$, the calculated value of ζ decreases, while Q decreases (see Discussion). However, as we use a single value of $F_{\rm min}$ in the entire analysis, this does not affect any differences seen between subsets of the data. We used the PYTHON/SCIPY function SCIPY.STATS.ANDERSON_KSAMP to compute the AD-test; as this function is only designed to convert the AD TS to p-values between $p = 0.001$ and 0.25, we supplemented this conversion with Monte Carlo simulations as detailed in Appendix A.

In the AD-test comparison between the synthetic and observed count rates, the time-dependence of these quantities is ignored. However, given that flares may last longer than the choosen time bin size (500 s), the count rates in consecutive time bins are not necessarily independent (Li et al. 2015). In our analysis, we will therefore observe a steeper flare flux distribution (i.e. a relatively higher number of faint flares), as the total flare fluence is divided over multiple time bins. This paper is mainly focused on the comparison between sub-sets of the *Swift* campaign, where the same steepening of the distribution is expected. Therefore, we argue that sub-sets of the data can accurately be compared with this method.

Before moving to the 1D method below, we make a brief note regarding the above method to calculate confidence intervals. In addition to the method described above, Neilsen et al. (2015) also applied an MCMC fitting routine to directly calculate and maximize the likelihood from the *Chandra* light curve. Comparing these two methods, one finds that both approaches return consistent results for the best-fitting parameters. However, the AD-test method returns slightly more conservative estimates of the confidence levels. Due to its computational speed, allowing us to perform a variety of tests on the long-term *Swift* data, we opted to use the AD-test method. However, aiming to remain conservative in our comparison between different sub-sets of the *Swift* data (see Section 3), we did not correct for a possible over-estimation of confidence levels.

2.2.2 One-dimensional method

The above method assumes that Q and ζ are independent, while not all combinations will be able to represent the real light curve: for instance, a high Q will require a small ζ to prevent systematically overpredicting the light-curve rates. Therefore, as a complementary 1D method, we again assumed each point in our light curve consists of the sum of a quiescent and a flaring component, both with Poisson noise. Therefore, for each

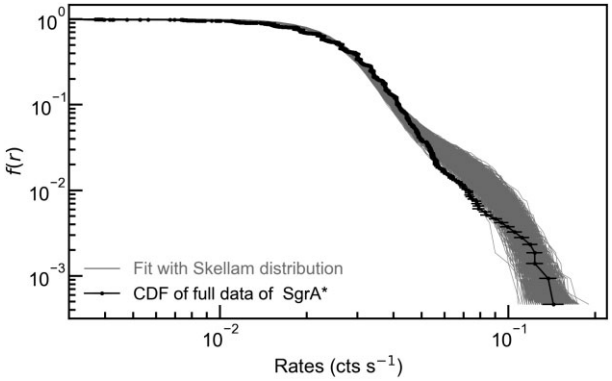

Figure 3. Empirical CDF of count rates of Sgr A* for the entire 2006–2019 *Swift* data set (as described in Section 2.1) is shown in black. The grey lines are 1000 sets of simulated data following a Skellam distribution. We note that in this figure, model components should be added vertically; likewise, the CDF uncertainties are vertical, as the fractions are set by the number of data points in the light curve.

data point, we can state that the rate r_i is given by $r_i = (1/\Delta t)\,[\mathtt{prand}(Q\Delta t) + \mathtt{prand}(F \times (r_{\max}/F_{\max})\Delta t)]$. Since the sum of two Poisson random variables with mean λ_1 and λ_2 is again a Poisson random variable with mean $\lambda_1 + \lambda_2$ and the fact that Δt is a constant, then the mean of all the observations in our data set will be the sum of the mean of both components: $\overline{r} = \overline{Q} + \overline{F} \times (r_{\max}/F_{\max})$.[5] Computing $\overline{F}$ from equation (2), we can express the mean count rate for our light curve as

$$\overline{r} = \overline{Q} + \frac{r_{\max}}{F_{\max}} \left(\frac{1 - 1/\zeta}{2 - 1/\zeta}\right) \frac{F_{\max}^{2-1/\zeta} - F_{\min}^{2-1/\zeta}}{F_{\max}^{1-1/\zeta} - F_{\min}^{1-1/\zeta}}, \tag{5}$$

where the second term of the right-hand side of equation (5) corresponds to mean value of fluxes, $\overline{F}$.

Therefore, we calculated the mean count rate of our data, $\overline{r}$, and we assumed fixed values of $F_{\max}$, $F_{\min}$, and $r_{\max}$. For a given ζ, we computed $\overline{F}$ and then we used equation (5) to calculate $\overline{Q}$. With those ζ and $\overline{Q}$, we generate 1000 sets of synthetic data, and stored the average TS from the AD-test. We repeated this process on a non-linear grid of ζ between 0 and 1 (focusing on values between 0.4 and 0.8 with higher resolution after inspection of the final PDF shapes for low-resolution grids). We then converted the average TS at each ζ into p-values and again we calculated the median value of ζ with its 1σ confidence interval.

This method has the advantage to be ∼100 times quicker than the 2D method, so it allowed us to run more simulations and re-draw data in order to find significance of changes in our analysis (see Section 3). The downside is that it assumes the mean count rate, $\overline{r}$, does not have an uncertainty; therefore, we also applied the 2D method to compare their results for the analysis of individual years.

3 RESULTS

The empirical CDF of Sgr A* obtained for the entire 2006–2019 light curve is shown in black in Fig. 3. We found that the quiescent component of Sgr A* emission is best represented by a pure Poisson distribution with mean count rate of $Q = 22.3^{+1.6}_{-1.4}$ count ks^{-1},

[5]Here, we ignore the background region count rate as we assume that the background is spatially constant. The introduction of the Skellam statistics is therefore only necessary when considering the distribution of subtracted data, not the mean.

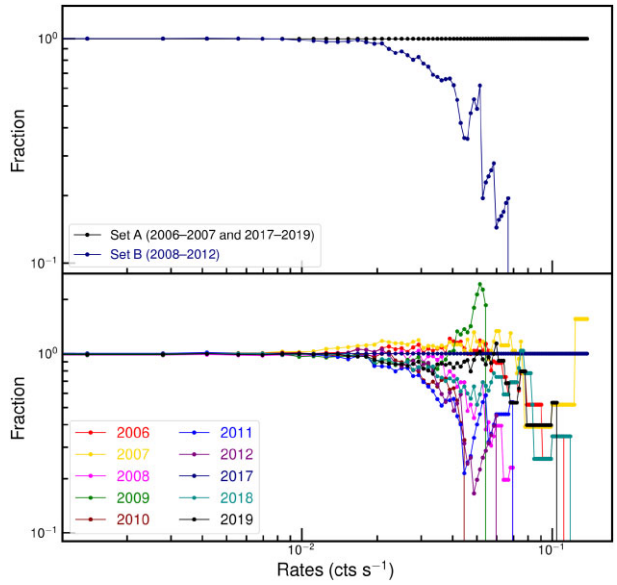

Figure 4. *Top panel:* fraction of the CDF for Set *A* (2006–2007 and 2017–2019) and Set *B* (2008–2012). The reference line in black is the fraction $f_A(r)/f_A(r)$. *Bottom panel:* Same as top panel, but the ratio is for individual years. The reference line in navy blue is the fraction $f_{2017}(r)/f_{2017}(r)$.

and a power-law process with $\zeta = 0.57^{+0.18}_{-0.12}$; for both parameters, throughout this paper, we will report the mode, with uncertainties corresponding to the 1σ level. These values of ζ and Q were calculated with the 1D method, while the 2D method returns consistent values. The grey lines in Fig. 3 are 1000 different synthetic sets, being the combination of both distributions as established in equation (4). We can notice there is a clear model excess towards the tail of the distribution, especially between ∼0.05 and 0.1 count s^{-1}, corresponding to the flaring component in the synthetic data. This discrepancy cannot be improved by changing the value of ζ nor Q in the simulations. It also implies that the full 2006–2019 *Swift* data cannot be described by a stationary Poisson + power-law model as well as the *Chandra* 2012 data in Neilsen et al. (2015), despite the higher sensitivity of *Chandra*.

Since Neilsen et al. (2015) showed that this model can reproduce the 2012 *Chandra* count rate distribution of Sgr A*, while it does not appear to work similarly well for all *Swift* data, time-variability in the CDF might play a role. To study whether this might be the case, we first divided the data into two sets and reanalysed the CDFs: Set *A*, containing the 5 yr with the highest maximum count rates measured by *Swift* (i.e, 2006–2007 and 2017–2019), and Set *B*, containing the 5 yr with the lowest maximum count rates (2008–2012). We only applied the 1D method for this, as testing the significance of any differences between the two sets required simulations that were too computationally expensive in the 2D method (see below).

The distinction between these sets can be seen in the top panel of Fig. 4, where we plot the ratio of the CDF for Sets *A* and *B*, both with respect to the CDF of Set *A*. The clear difference visible between the CDFs from Sets *A* and *B* indeed hints towards changes in flaring rate over time: the CDFs of the two sets start to diverge around ∼0.02 count s^{-1}, much lower than the flare rates used to select Sets A and B. Therefore, we first tested, using the AD-test, whether Sets A and B are consistent with being drawn from the same distribution. We find $TS = 3.30$, corresponding to a p-value of 0.015. The CDF modelling for these two sets is shown in the top left- and top right-hand panels of Fig. 5: We can notice a visual improvement in these

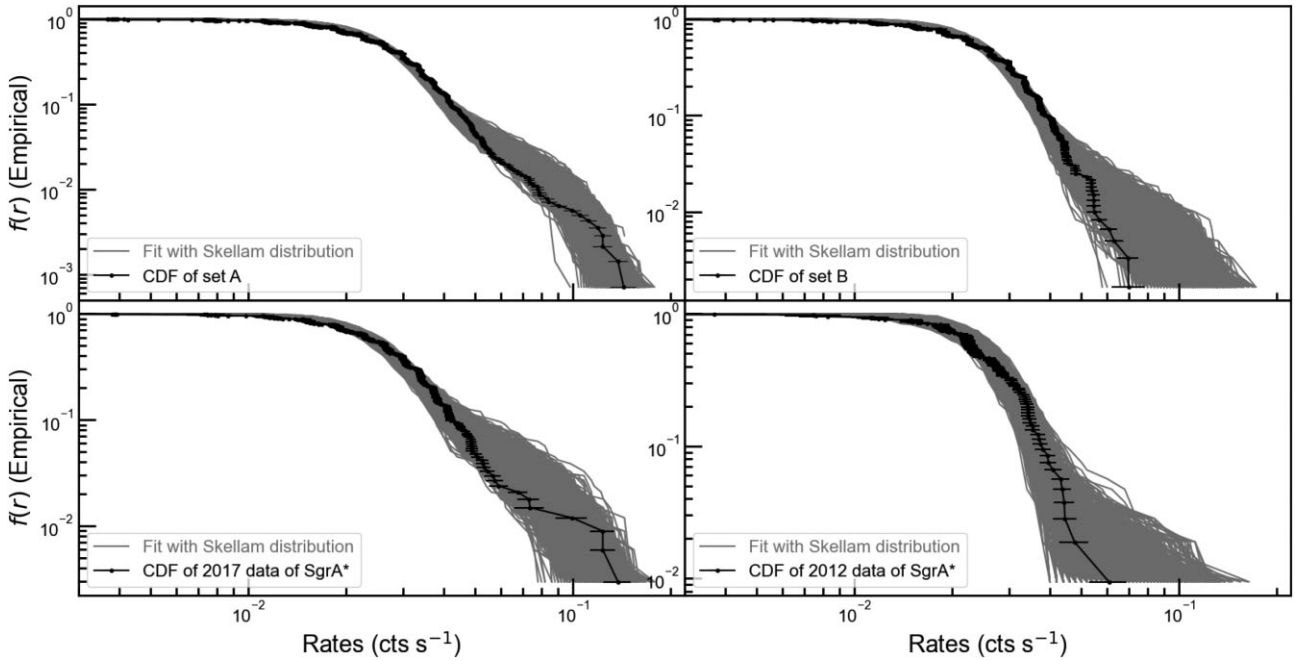

Figure 5. Results of the fits for: Top left-hand panel: Set *A*, consisting of the years with the highest count rates of Sgr A*; Top right-hand panel: Set *B*, consisting of the years with the lowest count rates of Sgr A*; Bottom left-hand panel 2017 data of Sgr A*; Bottom right-hand panel: 2012 data of Sgr A*. All these fits correspond to the values of ζ and Q calculated with the 1D method.

Table 1. Measured values for the quiescent (Q) and flaring (ζ) parameters found for different data sets of Sgr A* data. As the best-fitting value, we report the mode of the distribution, while uncertainties are quoted at 1σ sigma and upper limits at 90 per cent. Set A consists of 2006, 2007, 2017, 2018, 2019, while the rest of the years comprise Set B.

Year	ζ (2D)	Q (2D) (count ks^{-1})	ζ (1D)	Q (1D) (count ks^{-1})	N
Set *A*			$0.59^{+0.04}_{-0.08}$	$22.1^{+2.0}_{-1.7}$	1406
Set *B*			$0.52^{+0.03}_{-0.09}$	$22.1^{+1.5}_{-1.2}$	603
2006	$0.65^{+0.08}_{-0.09}$	$21.4^{+1.6}_{-1.7}$	$0.59^{+0.03}_{-0.08}$	$23.2^{+1.9}_{-1.6}$	323
2007	<0.71	$19.5^{+2.6}_{-1.4}$	$0.58^{+0.03}_{-0.12}$	$24.9^{+2.4}_{-1.9}$	215
2008	$0.64^{+0.07}_{-0.21}$	$22.4^{+1.6}_{-3.0}$	<0.57	$23.7^{+1.6}_{-2.3}$	242
2009	<0.85	$22.4^{+4.5}_{-2.5}$	<0.70	$22.0^{+3.8}_{-4.4}$	49
2010	<0.70	$20.6^{+3.5}_{-1.1}$	<0.54	$23.0^{+1.8}_{-2.2}$	102
2011	<0.73	$21.8^{+4.0}_{-1.0}$	<0.58	$21.8^{+2.0}_{-2.6}$	104
2012	<0.68	$19.8^{+3.3}_{-1.0}$	<0.53	$23.9^{+1.5}_{-2.2}$	106
2017	$0.69^{+0.06}_{-0.09}$	$23.1^{+1.5}_{-1.7}$	$0.62^{+0.04}_{-0.05}$	$21.5^{+1.7}_{-1.7}$	335
2018	$0.65^{+0.06}_{-0.09}$	$24.2^{+1.3}_{-1.7}$	$0.58^{+0.03}_{-0.06}$	$20.7^{+1.8}_{-1.6}$	323
2019	$0.67^{+0.08}_{-0.11}$	$24.5.5^{+1.6}_{-2.1}$	$0.61^{+0.03}_{-0.07}$	$20.3^{+2.0}_{-1.8}$	210

fits compared with the full set, particularly for Set *A*; In Set *B*, the flaring rate appears to be overestimated still in a majority of plotted synthetic light curves. The values of ζ and Q that we measured for these sets are summarized in Table 1.

The possible presence of a difference in the flaring component between Sets *A* and *B*, translates into a difference in fitted power-law index, although both values are consistent within their 1σ intervals – see also their probability density functions of ζ in the top panel of Fig. 6. We measured the difference in power-law index with the parameter $\Delta\zeta = \zeta_A - \zeta_B = 0.07$. However, this difference and the

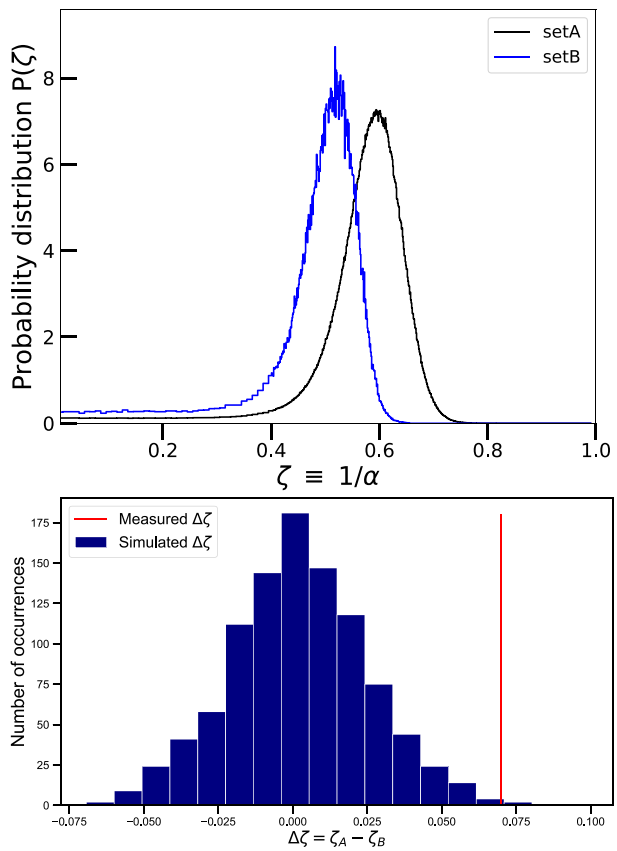

Figure 6. Top panel: probability distribution as function of ζ for Sets A and B. The maximum in the graphs gives the best ζ for each set. Bottom panel: histogram of occurrences of $\Delta\zeta$ in randomized data of Sets *A* and *B* using a fix value of $F_{\rm min} = 0.10$. The measured value of $\Delta\zeta$ is shown in red. Most of the simulations have a $\Delta\zeta \sim 0.0$, while only three simulations show $|\Delta\zeta| > 0.07$.

apparently better fits for the two sets compared to the complete light curve, might simply be due to the smaller number of data points. Therefore, we tested the hypothesis that such a change could arise at random when splitting the data in two, in the following way: we created two sets with the same number of data points of Sets *A* and *B*, but the count rates were taken randomly from the entire set of data. We then fitted those sets and calculated the value of $\Delta\zeta$ and repeated the process 1000 times. This test was performed by applying the 1D method, taking advantage of its faster performance. The histogram of occurrences for $\Delta\zeta$ is shown in the bottom panel of Fig. 6. We can observe that $|\Delta\zeta| \geq 0.07$ occurs merely three times in the randomized data. Therefore, the probability that a change in flaring distribution of this magnitude arises by chance, *only* due to the decrease in data points, is ~ 0.3 per cent. We have repeated this test for different values of $F_{\rm min}$ ($F_{\rm min} = \{0.05, 0.1, 0.2\}$), redoing both, the fit of the real data and creating 1000 synthetic data sets; the same conclusion can be drawn in those cases.

With these results, we decided to further investigate the possible variability by dividing the data into individual years. First, we used a 1-sample AD-test, to test whether the data for the individual years is inconsistent with a pure Poisson process. We find that for the years 2010, 2011, and 2012, we cannot reject the hypothesis that a pure Poisson process underlies their light curve at $p < 0.01$: their p-values are 0.03, 0.01, and 0.05, respectively. For the other years, we do find $p < 0.01$, indicating the significant presence of a flaring component.

Then, we turned to analysing the CDFs of the individual years. The results are summarized in Table 1, and the corresponding contour plots, probability distributions and CDFs fits are presented in Fig. B1. We find that for several years, only an upper limit on the flaring parameter ζ can be measured. In those cases, the (marginalized) probability density function of ζ does not tend to zero as $\zeta \rightarrow 0$, and we list the 90 per cent upper limit on the ζ instead of the mode. We observe this effect in 2009–2012 for both methods and 2008 in the 1D method only. In the 2007 data, an upper limit is obtained from the 2D method, while for the 1D method, the probability as $\zeta \rightarrow 0$ is close to, but not exactly, zero. For this borderline case, we therefore list the mode with uncertainties in Table 1, but note that its 90 per cent upper limit would be $\zeta < 0.63$. We can see this comparison of individual years graphically in Fig. 7: The top panel shows how the ζ-Q contour plots from the 2D method do not close for the 2012 data, while they do for the 2017 data. While overlapping, the probability contour of 2012 data is systematically shifted to the left compared with the one for 2017 data. In the bottom panel of Fig. 7, we show the probability density, $p(\zeta)$ from the 1D method as function of ζ. A clear maximum is again observed for 2017 data, but not for 2012 data, where after reaching the maximum, $p(\zeta)$ remains almost constant, indicating that for lower ζ, fits of similar quality are obtained. It is for those years that show similar behaviour, that we calculate a 90 per cent upper limit on ζ (in no cases do we find that the resulting probability interval of Q is unclosed).

Splitting up the data, one can wonder whether the inability to constrain ζ in several years is simply due to a lower observing cadence: from 2009–2012, the *Swift* cadence was only one observation per 3 d, and in 2009, there were no observations between February and April (Degenaar et al. 2013). To test the effect of a lower cadence, we simulated synthetic data sets with a ζ and Q similar to those found in 2006–2007 or 2017–2019, but with the low cadence of the years between 2009–2012. When we reapplied the 2D method to these synthetic low-cadence years, we do find closed contour plots for ζ between 0.01 and 1.00 with the 2D method. A second test is to look more closely at Table 1: notably, in 2019 contained a smaller number of data points (210) than 2007 (215)

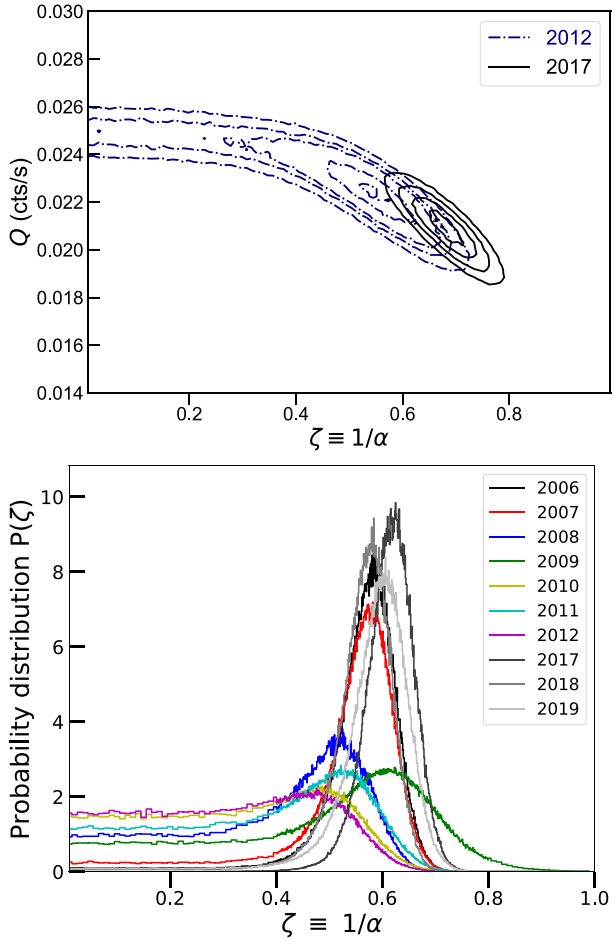

Figure 7. *Top panel:* probability contour plots in the $\zeta - Q$ plane for 2017 data (black lines), 2012 data (blue dashes). *Bottom panel:* probability density from the 1D method, as function of ζ for 2017 and 2012 data.

and 2008 (242); however, in 2019, ζ can clearly be measured, while only an upper limit could be measured for 2008 (1D-method) and 2007 (2D-method), and 2007 is also a borderline case in the 1D method.

Therefore, we conclude that a lower cadence can contribute to poorer constraints on ζ, for instance, in 2009 (with only 49 points), but cannot account for all differences between years. This behaviour, together with the inability of a single flare distribution to reproduce all 2006–2019 observations, and the unconstrained ζ parameter between 2009 and 2012, can be interpreted as additional evidence for a change in the X-ray flaring emission properties from Sgr A*. We show all other contour maps, CDF fits, and probability density functions of ζ for individual years in Appendix B.

4 DISCUSSION

The fits presented in this work of the CDF of Sgr A* count rates measured with *Swift*, show that it is not possible to adequately describe the complete set of X-ray observations using the model of a single power-law process combined with a pure Poisson process, as shown in Fig. 3. The excess in the tail of the distribution is not due to the power-law index only, but also due to the high Poisson mean rate used in the quiescent component. Creating synthetic data sets with less flaring activity (i.e. lower ζ) would require a higher value of Q in the quiescent component, and then the lower count rates will be overes-

timated. Nevertheless, the emission can be described with this model better once we divide the data in different sets, as shown in Fig. 4.

Looking at the values from the 1D method in Table 1, we notice that the flaring properties of the years forming Set *A* appear similar: the difference in the best-fitting ζ is small in those years, ranging from 0.62 to 0.58. On the other hand, for the years forming Set *B*, the 1D-approach returns similar upper limits on ζ, with the exception of 2009; these upper limits are typically below the best-fitting flaring parameters for the years in Set *A*, but do overlap with their uncertainty intervals. This result may suggest that the issues in fitting the full set of *Swift* data are not necessarily due to an incorrect model – as expected given the adequate description of the *Chandra* XVP data with this model (Neilsen et al. 2015). Instead, it could also fit with the idea that the flaring rate of Sgr A* has changed on the time-scales of years, while the first model fit assumes it remains constant over the considered time range.

A change in flaring properties on years time-scales is further supported by other lines of reasoning: the consistency of the light curves in 2010, 2011, and 2012 with a pure Poisson process; the non-zero probabilities for low ζ in the individual years making up Set B, which cannot be fully explained by their low cadence alone (e.g. comparing 2019 with 2007 and 2008); and the shift in the probability density function of ζ between Sets *A* and *B*. We note that the best-fitted ζ values for those two sets do not change beyond the 1σ confidence level. However, combined, the lines of evidence above are consistent with a decrease in flaring activity between 2008 and 2012.

Before discussing such variations in the physical picture of Sgr A*, we briefly turn to the assumptions of the method. First, the value of $F_{\rm min}$ also affects the fits and measured parameters in the data. Regarding the 1D method, in the top panel of Fig. 8, we show again $p(\zeta)$ for the 2017 data, assuming multiple values of $F_{\rm min}$. The ζ that best match the observations gives the maximum in each graph. It is noticeable that a higher $F_{\rm min}$ returns a lower ζ. From equation (5), we see that as the value of ζ decreases, it reduces the value of $\overline{Q}$, and in fact, the 1D method breaks for any $F_{\rm min} > 0.30$, as this would require negative values of $\overline{Q}$ from equation (5). The same effect, albeit at different exact values of ζ is observed in the analysis of any of the other years. Moreover, this also can be observed when the 2D method is applied. Since our study compares subsets of the *Swift* monitoring using the same $F_{\rm min}$, the dependence of ζ on $F_{\rm min}$ does not affect our qualitative conclusions. However, Fig. 8 does show that the absolute value of ζ found in the analysis of a single data set cannot be taken as a unique description of the underlying flare flux distribution. The best-fitting values in Table 1 should therefore only be interpreted in relation to each other, preferably comparing the full distributions shown in Appendix B.

Not only $F_{\rm min}$ affects the measured values of Q and ζ; comparing the columns in Table 1 shows that the 1D and 2D approaches also find slightly different best-fitting parameters. The best-fitting parameters are, however, consistent within their 1σ uncertainties. Importantly, we find that the 1D method is better able to constrain the ζ-parameter. A possible reason for the difference in the exact best-fitting value might be that the 1D method ignores the uncertainty on the measured count rate: this uncertainty turns the line probed by the 1D method in the ζ–Q plane, into a band, which is automatically fully explored in the 2D method.

We also notice that dividing the data into different sets, i.e. having smaller number of data points in each, makes the fits statistically better (i.e. reaching lower TS values). This happens also when we randomized the data for individual years, as described in Section 3 (testing the differences between Sets A and B). We show this more explicitly in Fig. 8 (bottom): in blue, we plot the histogram of TS

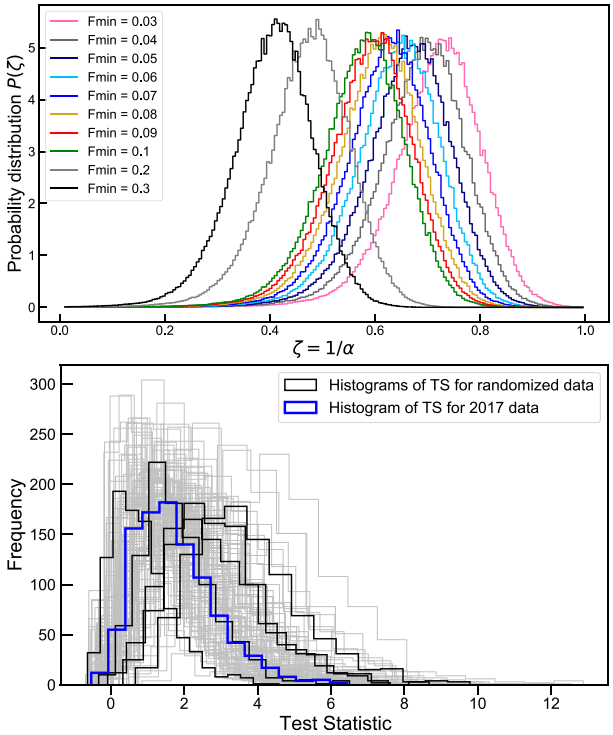

Figure 8. Top panel: Probability density as function of ζ using multiple values of $F_{\rm min}$ with 2017 data of Sgr A*. For a given $F_{\rm min}$, the best ζ corresponds to the maximum in the graph. Bottom panel: Distribution of TS values for the best-fitting ζ and Q for 2017 data (blue) and for 100 randomized data sets with the same number of data points.

values for the comparisons between the 2017 data and the 1000 synthetic light curves generated using the best-fitting values of ζ and Q (see Table 1). In the same figure, in grey, we plot the same histograms, for 100 light curves of the 2017 length, randomly picked from the full data set; we also plot a small number these in black, to highlight their shape. The similarity in the distributions indicates that the statistical quality of the *best* fits are similar, confirming that smaller light curves are easier to fit satisfactorily.

However, we stress that the above analysis only compares the TS values for the best-fitting parameters: whether changes between subsets of the data are found, depends also on whether a satisfactory fit is only found for a limited range of parameters. This comparison therefore does not imply that the fits are *only* better due to the smaller number of data points, as discussed in Section 3 as well: even at low-cadence, synthetic light curves created using the best-fitting parameters of 2006–2007 and 2017–2019 do not return ζ upper limits, and the comparison between 2008, 2009, and 2019, yields a similar suggestion. However, we cannot exclude the possibility that the lower cadence contributes to the better fits.

Turning to the interpretation of the possibly variable flaring rate, we can first compare our work with the results from Ponti et al. (2015), which reports an increase in Sgr A* emission at the end of 2013 and 2014. Our work shows evidence that Sgr A*'s emission in the X-ray band presents different properties between years, and although we are ignoring 2013–2016 data, the activity of Sgr A* in the last years of observation of *Swift* shows higher count rates than the rest of the years. On the other hand, work performed by Yuan & Wang (2015) and Bouffard et al. (2019) reported no sign of changes in Sgr A*'s activity using *Chandra* observations during different epochs. Possibly the flaring rate of Sgr A* doesnot show measurable

changes on time-scales shorter than a year, as shown by the work from Neilsen et al. (2015). This would explain why such short-time-scale changes are not seen by either *Swift* or *Chandra*, while they may occur on few-years time-scales. Hence, we need multiyear, recurring monitoring to further test and confirm any of these results.

The nIR emission from Sgr A* also shows strong flaring variability, with fluxes that have been modelled as either a power-law or lognormal distribution (Witzel et al. 2012, 2018) with an additional high-flux tail (Dodds-Eden et al. 2011), or a log-lognormal distribution (Meyer et al. 2014). Investigating the emission mechanism of the nIR and X-ray flares, Neilsen et al. (2015) assume that the X-ray and nIR fluxes both follow power-law distributions and that these fluxes are coupled. These simple assumptions imply a direct relation between the power-law indices in nIR and X-rays. In that scenario, the observed variability in X-ray flaring rate would be accompanied with similar changes in the nIR flaring rate. However, if such changes in the nIR are not present, the nIR–X-ray coupling either changes over time, or the X-ray and nIR flares are not (always) related. If the X-ray and nIR flares are indeed manifestations of the same underlying emission process, we deem it unlikely that the coupling between wavelengths varies over years time-scales and instead expect a similar increase in nIR flaring rate as the X-ray activity increases.

The nIR flux distribution of Sgr A* indeed varies on time-scales of years; for instance, Dodds-Eden et al. (2011) and Witzel et al. (2012) both analyse overlapping, but not identical, sets of *VLT/NACO* observations and report different flare flux distributions (both in shape and parameters). More recently, analysing *Keck* nIR observations obtained in 2019 April and May, Do et al. (2019) reported an increase in flaring rate compared to *Keck* observations between 2005 and 2013. The increased 2019 flaring rate coincides with the increased flaring rate seen in our *Swift* monitoring, while during the 2005–2013 interval, the *Swift* monitoring is dominated by low flaring rate years (2008–2012, compared to 2006–2007 with a higher flaring rate). These comparisons are suggestive of a coupled change in flaring rate between nIR and X-rays, which could be confirmed by a time-resolved analysis of the nIR and X-ray fluxes on year time-scales in future data.

Semi-analytical studies generally favour synchrotron emission from a non-thermal population of electrons as the source of simultaneous nIR/X-ray flaring (e.g. Dibi et al. 2016). The cooling of synchrotron electrons in these models results in the steepening of the spectral slope from the nIR to the X-rays and therefore, provides a crucial link between the two flare populations (e.g. Dodds-Eden et al. 2010; Dibi et al. 2014; Ponti et al. 2017). Magnetized blobs of plasma that naturally form in the turbulent accretion flow (e.g. Sironi, Petropoulou & Giannios 2015; Ripperda et al. 2021) are said to be a viable source of rapid flaring activity in Sgr A* (e.g. Ball et al. 2016; Chatterjee et al. 2021; Scepi, Dexter & Begelman 2021). Indeed, Gutiérrez, Nemmen & Cafardo (2020) suggests that a large enough magnetized blob could produce the high levels of non-thermal synchrotron emission that would be required to explain the Do et al. (2019) nIR flare.[6] Furthermore, an increase in the accretion rate could result in a strongly magnetized accretion disc as more and more magnetic fields are advected in. Such a disc is able to produce magnetized blobs with relativistic temperatures (Ripperda, Bacchini & Philippov 2020; Ripperda et al. 2021) that leads to stronger flares. Indeed, an increase in the X-ray flare rate in the period of 2017–2019 does seem to support this possibility. Current numerical models of accreting black holes are only able to address a few 10 s of hours of Sgr A* activity (e.g. Chan et al. 2015; Ball et al. 2016; Chael et al. 2018; Chatterjee et al. 2020), and much longer simulations are necessary to predict whether a small change in the amount of accreted material on the time-scale of years could trigger a large non-thermal event close to the black hole.

Do et al. (2019) also suggest an alternative explanation for the observed increase in nIR activity: they suggest it could be related to the periastron passage in 2018 May of the windy star S0-2. However, if the increase in nIR and X-ray activity are indeed coupled, this explanation is inconsistent with the increased X-ray activity already observed in 2017 (see also Ressler, Quataert & Stone 2018). Alternatively, the increase in activity might result from the periastron passage of the gaseous object G2 in 2014 (Gillessen et al. 2012, 2013a,b; Madigan & McCourt 2016). Simulations of the response of Sgr A* to this periastron passage predicts a delay in increased activity of a few to ∼10 yr (Schartmann et al. 2012; Kawashima, Matsumoto & Matsumoto 2017). Without information from *Swift* between 2012 and 2017, we can infer an upper limit on the possible X-ray response to the G2 passage of ∼3 yr, lasting for 2 yr at the time of writing, thus fitting with those model predictions.

Swift monitoring also suggests a higher flaring rate before 2008, followed by a low-activity period between 2008 and 2012. Therefore, if the G2 object is responsible for the current high levels of activity, a similar object could have passed close by Sgr A* during the late 1990s or early 2000s to explain earlier enhanced activity. Indeed, a gaseous object similar to G2 (Clénet et al. 2004a,b, 2005; Ghez et al. 2005), named G1 by Pfuhl et al. (2015), passed similarly close to Sgr A* in 2001 (Witzel et al. 2017). We stress that these considerations are highly speculative; for instance, the *Swift* monitoring is only consistent with an increased flaring rate, but not with a change in the steady accretion activity of Sgr A*, although this might be hidden in the diffuse X-ray emission unassociated with Sgr A* due to *Swift*'s point spread function. Also, Yuan & Wang (2015) did not find evidence for changes in X-ray activity observed by *Chandra* after the passage of G2. However, if the enhanced activity is related to passages of objects such as G1 and G2, the decrease of flaring activity in 2008 – roughly 7 yr after G1's periastron passage – would predict an end to the current high level of flaring activity around ∼2021.

With continued monitoring in X-rays (by *Swift*) and nIR, the above prediction could be tested, although we note the importance of a regular and high cadence of these observations, similar to the 2017–2019 data, in order to eliminate any possible effects of sample size. Alternatively, new *Chandra* observations could similarly help to assess the variability on years time-scales, allowing for a comparison with the 2012 campaign. Finally, we remind the reader how this work focused on the broad statistical properties of the data set, and did not analyse individual flares; detailed statistical flare searches in the long-term *Swift* data, such as performed by Ponti et al. (2015), could be an alternative route to explore this Sgr A* data set.

5 CONCLUSIONS

We have analysed the empirical CDF of the X-ray emission from Sgr A* using *Swift* data for 2006–2012 and 2017–2019. By assuming the X-ray emission of Sgr A* is composed of a quiescent and a flaring component, we modelled it as the sum of a constant flux and a power-law flux distribution, respectively. We adopted two different methods to fit the CDF, yielding consistent results: a 2D method allowing the flare parameter ζ and the quiescent parameter Q to vary independently, and a 1D method establishing a dependence between ζ and Q as shown in equation (5). We found that the full data set cannot be correctly described with this model, while individual years do follow it better.

[6]We note that the flare visible in May 2019 in *Swift* monitoring (Fig. 1) did not occur simultaneously with the Do et al. (2019) nIR flare.

More concretely, we found that the data can be divided in two different groups: (1) a set of data with high flaring rate (2006–2007 and 2017–2019), which we called Set *A*, and (2) a set of data with low flaring rate (2008–2012), labelled as Set *B*. The main difference between these two sets can be summarized as:

(i) All years forming Set *A* have a well-constrained value of ζ, as shown in the bottom panel of Fig. 7. Therefore, the model of constant + power-law distribution can describe the data of these years. However, the years of Set *B* present larger errors on ζ or it cannot be constrained and only an upper limit can be established. In some cases in Set *B*, the emission can be well described by a pure constant model with no flaring at all.

(ii) The values of ζ for Set *B* may be systematically shifted to lower values than the ones obtained for Set *A*. Such lower values of ζ would indicate a lower flaring rate for Set *B*. While we highlight that these changes are within the 1σ confidence level, they may hint towards a change in the flaring rate between different years.

Finally, we tested explicitly the effect of dividing the data into two or more sets and we found that the probability of the improvement in the fits is due by chance when reducing the number of data points is just ~ 0.3 per cent. Similarly, the poorly constrained parameters for some of the years in Set *B* cannot be explained by only the lower number of data points of those sets. Therefore, we interpret these results as evidence for a change in the flaring rate of Sgr A*, being more active during 2006–2007 and 2017–2019, compared with the years of 2008–2012.

ACKNOWLEDGEMENTS

The authors thank the referee for an insightful and constructive report that improved the quality of this work. AA, JvdE, and ND were supported via an NWO/Vidi grant awarded to ND. AA gratefully acknowledges support from the ASPIRE programme and thanks the University of Amsterdam for hospitality. JvdE is also supported by a Lee Hysan Junior Research Fellowship from St Hilda's College, Oxford. KC and SM are supported by the NWO/VICI grant (no. 639.043.513) awarded to SM. KC is additionally supported by a Black Hole Initiative Research Fellowship from Harvard University, funded by the Gordon and Betty More Foundation, John Templeton Foundation, and the Black Hole PIRE programme. GP acknowledges funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 865637). COH is supported by NSERC grant RGPIN-2016-04602. DA acknowledges support from the Royal Society. This work made use of public data from the *Swift* data archive and data supplied by the UK *Swift* Science Data Center at the University of Leicester.

DATA AVAILABILITY

A data reproduction package for this paper can be found at the following DOI: 10.5281/zenodo.5140005

REFERENCES

Abuter R. et al., 2018, A&A, 618, L10
Anderson T. W., Darling D. A., 1954, J. Am. Stat. Assoc., 49, 765
Baganoff F. et al., 2001, Nature, 413, 45
Ball D., Özel F., Psaltis D., Chan C.-k., 2016, ApJ, 826, 77
Bouffard É., Haggard D., Nowak M. A., Neilsen J., Markoff S., Baganoff F. K., 2019, ApJ, 884, 148
Čadež A., Calvani M., Kostić U., 2008, A&A, 487, 527
Chael A., Rowan M., Narayan R., Johnson M., Sironi L., 2018, MNRAS, 478, 5209
Chan C.-k., Psaltis D., Özel F., Medeiros L., Marrone D., Sadowski A., Narayan R., 2015, ApJ, 812, 103
Chatterjee K. et al., 2020, MNRAS, 507, 5281
Chatterjee K. et al., 2021, MNRAS, 507, 5281
Chen Z. et al., 2019, ApJ, 882, L28
Clénet Y. et al., 2004a, A&A, 424, L21
Clénet Y. et al., 2004b, A&A, 417, L15
Clénet Y., Rouan D., Gratadour D., Marco O., Léna P., Ageorges N., Gendron E., 2005, A&A, 439, L9
Coti Zelati F. et al., 2017, MNRAS, 471, 1819
Degenaar N., Miller J., Kennea J., Gehrels N., Reynolds M., Wijnands R., 2013, ApJ, 769, 155
Degenaar N., Wijnands R., Miller J. M., Reynolds M. T., Kennea J., Gehrels N., 2015, J. High Energy Astrophys., 7, 137
Degenaar N., Reynolds M. T., Wijnands R., Miller J. M., Kennea J. A., 2019, Astron. Telegram, 12768, 1
Dibi S., Markoff S., Belmont R., Malzac J., Barrière N. M., Tomsick J. A., 2014, MNRAS, 441, 1005
Dibi S., Markoff S., Belmont R., Malzac J., Neilsen J., Witzel G., 2016, MNRAS, 461, 552
Do T. et al., 2019, ApJ, 882, L27
Dodds-Eden K., Sharma P., Quataert E., Genzel R., Gillessen S., Eisenhauer F., Porquet D., 2010, ApJ, 725, 450
Dodds-Eden K. et al., 2011, ApJ, 728, 37
Eckart A. et al., 2004, A&A, 427, 1
Evans P. et al., 2007, A&A, 469, 379
Evans P. A. et al., 2009, MNRAS, 397, 1177
Gehrels N. et al., 2004, ApJ, 611, 1005
Genzel R., Eisenhauer F., Gillessen S., 2010, Rev. Mod. Phys., 82, 3121
Ghez A. M. et al., 2005, ApJ, 635, 1087
Ghez A. M. et al., 2008, ApJ, 689, 1044–1062
Gillessen S. et al., 2012, Nature, 481, 51
Gillessen S. et al., 2013a, ApJ, 763, 78
Gillessen S. et al., 2013b, ApJ, 774, 44
Goldwurm A., Brion E., Goldoni P., Ferrando P., Daigne F., Decourchelle A., Warwick R. S., Predehl P., 2003, ApJ, 584, 751–757
Gutiérrez E. M., Nemmen R., Cafardo F., 2020, ApJ, 891, L36
Haggard D. et al., 2019, ApJ, 886, 96
Kawashima T., Matsumoto Y., Matsumoto R., 2017, PASJ, 69, 43
Kennea J. A., Burrows D. N., Campana S., Godet O., Nousek J., Gehrels N., 2006, Astron. Telegram, 753, 1
Kennea J. A. et al., 2013, ApJ, 770, L24
Li Y.-P. et al., 2015, ApJ, 810, 19
Liu S., Melia F., 2002, ApJ, 566, L77
Liu S., Petrosian V., Melia F., 2004, ApJ, 611, L101
Madigan A.-M., McCourt M., 2016, MNRAS, 457, L89
Markoff S., 2010, Proc. Natl. Acad. Sci., 107, 7196
Markoff S., Falcke H., Yuan F., Biermann P. L., 2001, A&A, 379, L13
Meyer L., Witzel G., Longstaff F. A., Ghez A. M., 2014, ApJ, 791, 24
Morris M. R., Meyer L., Ghez A. M., 2012, Res. Astron. Astrophys., 12, 995
Neilsen J. et al., 2013, ApJ, 774, 42
Neilsen J. et al., 2015, ApJ, 799, 199
Nowak M. A. et al., 2012, ApJ, 759, 95
Pettitt A. N., 1976, Biometrika, 63, 161
Pfuhl O. et al., 2015, ApJ, 798, 111
Ponti G. et al., 2015, MNRAS, 454, 1525
Ponti G. et al., 2016, MNRAS, 461, 2688
Ponti G. et al., 2017, MNRAS, 468, 2447
Reid M. J., Brunthaler A., 2004, ApJ, 616, 872
Ressler S. M., Quataert E., Stone J. M., 2018, MNRAS, 478, 3544
Reynolds M., Kennea J., Degenaar N., Wijnand s R., Miller J., 2016, Astron. Telegram, 8649, 1
Reynolds M., Kennea J., Miller J., Degenaar N., Wijnands R., 2018, Astron. Telegram, 11313, 1
Ripperda B., Bacchini F., Philippov A., 2020, ApJ, 900, 14
Ripperda B., Liska M., Chatterjee K., Musoke G., Philippov A. A., Markoff S. B., Tchekhovskoy A., Younsi Z., 2021, ApJ, preprint (arXiv:2109.15115)

Scepi N., Dexter J., Begelman M. C., 2021, MNRAS, preprint (arXiv:2107.08056)
Schartmann M., Burkert A., Alig C., Gillessen S., Genzel R., Eisenhauer F., Fritz T. K., 2012, ApJ, 755, 155
Scholz F. W., Stephens M. A., 1987, J. Am. Stat. Assoc., 82, 918
Sironi L., Petropoulou M., Giannios D., 2015, MNRAS, 450, 183
Trap G. et al., 2011, A&A, 528, A140
Witzel G. et al., 2012, ApJS, 203, 18
Witzel G. et al., 2014, ApJ, 796, L8
Witzel G. et al., 2017, ApJ, 847, 80
Witzel G. et al., 2018, ApJ, 863, 15
Yuan Q., Wang Q. D., 2015, MNRAS, 456, 1438–1450
Yuan F., Quataert E., Narayan R., 2003, ApJ, 598, 301
Yuan Q., Wang Q. D., Liu S., Wu K., 2017, MNRAS, 473, 306–316
Yusef-Zadeh F., Wardle M., Heinke C., Dowell C. D., Roberts D., Baganoff F. K., Cotton W., 2008, ApJ, 682, 361

APPENDIX A: ANDERSON-DARLING TEST STATISTICS AND SIGNIFICANCE LEVELS

In order to test whether the generated synthetic light curve for a given parameter pair (ζ, Q) and the observed *Swift*/XRT light curve are consistent with the same underlying flux distribution, we employ the AD-test (Anderson & Darling 1954). As noted by Neilsen et al. (2015) in their work on the *Chandra* XVP data, and in earlier work by Scholz & Stephens (1987), this test is more sensitive for deviations in the tails of the distributions than for instance the Kolmogorov–Smirnov test, making it more suitable for comparing flaring behaviour on top of a second, quiescent component. However, a downside to the use of the AD-test is the lack of a single analytic expression relating the observed test statistic, referred to as TS in the main paper and A^2 in statistical literature, to a p-value: the probability that the two populations arise from the same underlying distribution.

Instead, early studies of the AD-test and its k-sample extensions calculated critical values for different sizes of the compared samples. In the two-sample case, these critical values are defined as the test statistic $A^2_{\rm crit}$, where $p(A^2_{nm} \geq A^2_{\rm crit}) < \alpha_{\rm crit}$, where α is the confidence level (we use the typical notation in statistical literature, not to be confused with the flaring parameter from Neilsen et al. 2015), and n and m are the sizes of the two compared samples. The critical values change as a function of the sample sizes and are often calculated for $\alpha_{\rm crit} = 0.01, 0.05, 0.10$ (Pettitt 1976; Scholz & Stephens 1987). Analytic extrapolations of these critical values can then be used to either find critical values for larger sample sizes n and m, or convert test statistics to probabilities for non-critical values. The PYTHON SCIPY function SCIPY.STATS.ANDERSON_KSAMP, which we used to calculate the AD-test statistic throughout this work, uses such extrapolations to return probabilities as well. However, these probabilities are capped at $p = 0.001$ and 0.25, limiting the use of this function in calculating AD probabilities.

An alternative route to convert the test statistics to probabilities is via Monte Carlo calculations. For two samples of sizes n and m, one can explicitly calculate the distribution of A^2_{nm} by calculating A^2_{nm} for all $(n + m)!/(n!m!)$ possible orders of all $(n + m)$ values. The probability associated with the measured test statistic $A^2_{\rm measured}$, i.e. $p(A^2_{nm} \geq A^2_{\rm measured})$, can then be calculated from the distribution. This approach has a significant limitation: the number of orderings $(n + m)!/(n!m!)$ quickly increases, yielding unfeasible computational times for our data sets: our sample sizes exceed 100 for all years except 2009, implying both n and m exceed 100, while $n = m = 100$ yields $\sim 10^{59}$ orderings. However, as Scholz & Stephens (1987) note, one can instead generate a large number of randomly picked orderings, and use their A^2_{nm}-values as an approximation of the full underlying distribution – the probabilities calculated from this Monte Carlo approach are an unbiased estimator of the real probability.

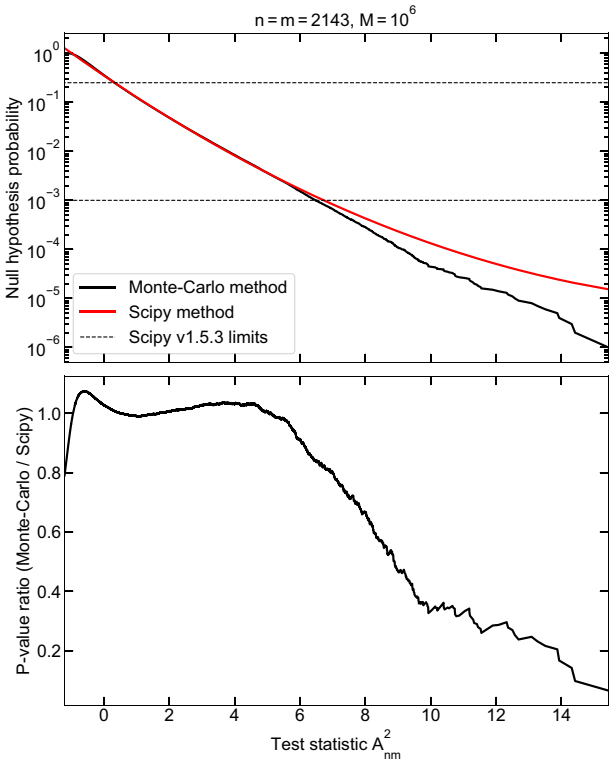

Figure A1. Top panel: Null hypothesis probability as function of AD-test statistic, using both SCIPY.STATS.ANDERSON_KSAMP v1.1.0 (red) and a Monte Carlo method (black). The dashed lines indicate the limiting p-values in SCIPY v.1.5.3. Bottom panel: the ratio between the probabilities calculated via the two methods plotted in the upper panel.

We therefore applied this Monte Carlo approach to convert measured test statistics to p-values. As the AD-test is a rank test, A^2_{nm} only depends on the order of the values of the two samples, and not on the values themselves. The null hypothesis in our analysis is that both samples are drawn from the same underlying distribution. Therefore, since A^2_{nm} does not depend on the values, we can define a list of integers ranging from 1 to $n + m$ as the underlying sample. In our analysis, n and m are equal, as we are comparing an observed light curve (for 1 y, set A/B, or the full data set) with a synthetic one of equal length. For each Monte Carlo iteration, we then create two samples, each of size n ($= m =$ half the full integer list), by randomly choosing integers from the list. We then calculate and save their A^2_{nm} using SCIPY.STATS.ANDERSON_KSAMP and repeat M times, where M is set by the intended accuracy in p-values.

To show the importance of using Monte Carlo calculations, we show a comparison between the test statistic to p-value conversion using SCIPY.STATS.ANDERSON_KSAMP and our approach in Fig. A1. We show the results for $M = 10^6$ iterations using $n = m = 2143$, which corresponds to the complete data set. The dashed lines indicate the minimum and maximum returned p-value using the most recent SCIPY (v1.5.3), while we used an older version (v1.1.0) without this restriction for comparison to show the effects at smaller and larger test statistics. Between the two limits, both approaches agree reasonably well, although the effect of applying an analytic interpolation is visible in the variable ratio in the lower panel. The discrepancy grows especially above test statistics of $\sim$5. As our method requires an accurate measurement of the p-values even for poorly fitting

parameters, in order to sample the entire parameter space, applying the Monte Carlo simulations is essential.

APPENDIX B: PARAMETERS ESTIMATION AND CDFS FOR INDIVIDUAL YEARS

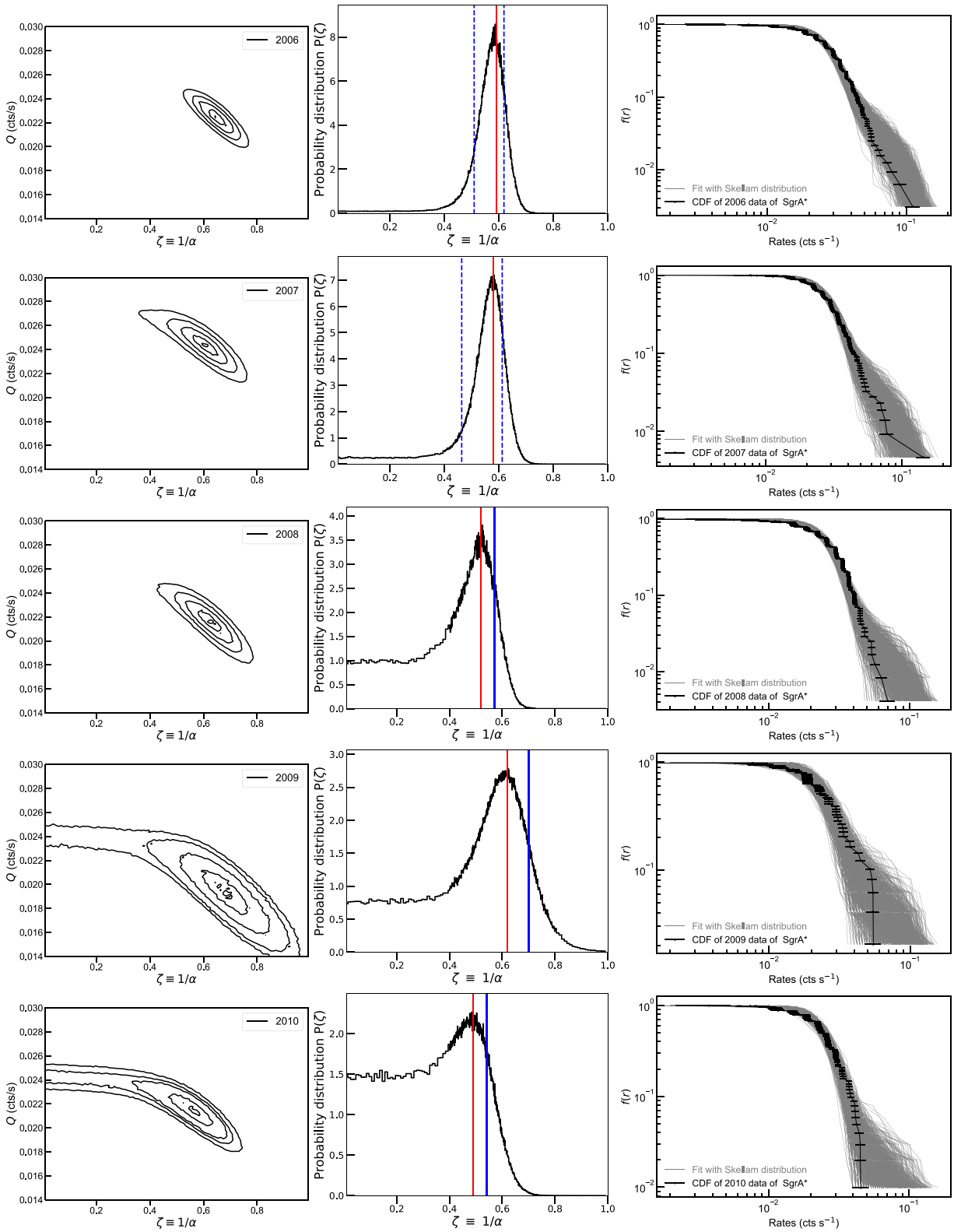

Figure B1. The contour plots in the $\zeta - Q$ plane for the 2D method and probability density function of ζ for the 1D method, for individual years, and fit of the corresponding CDF. Figures for years 2012 and 2017 are shown in the main body.

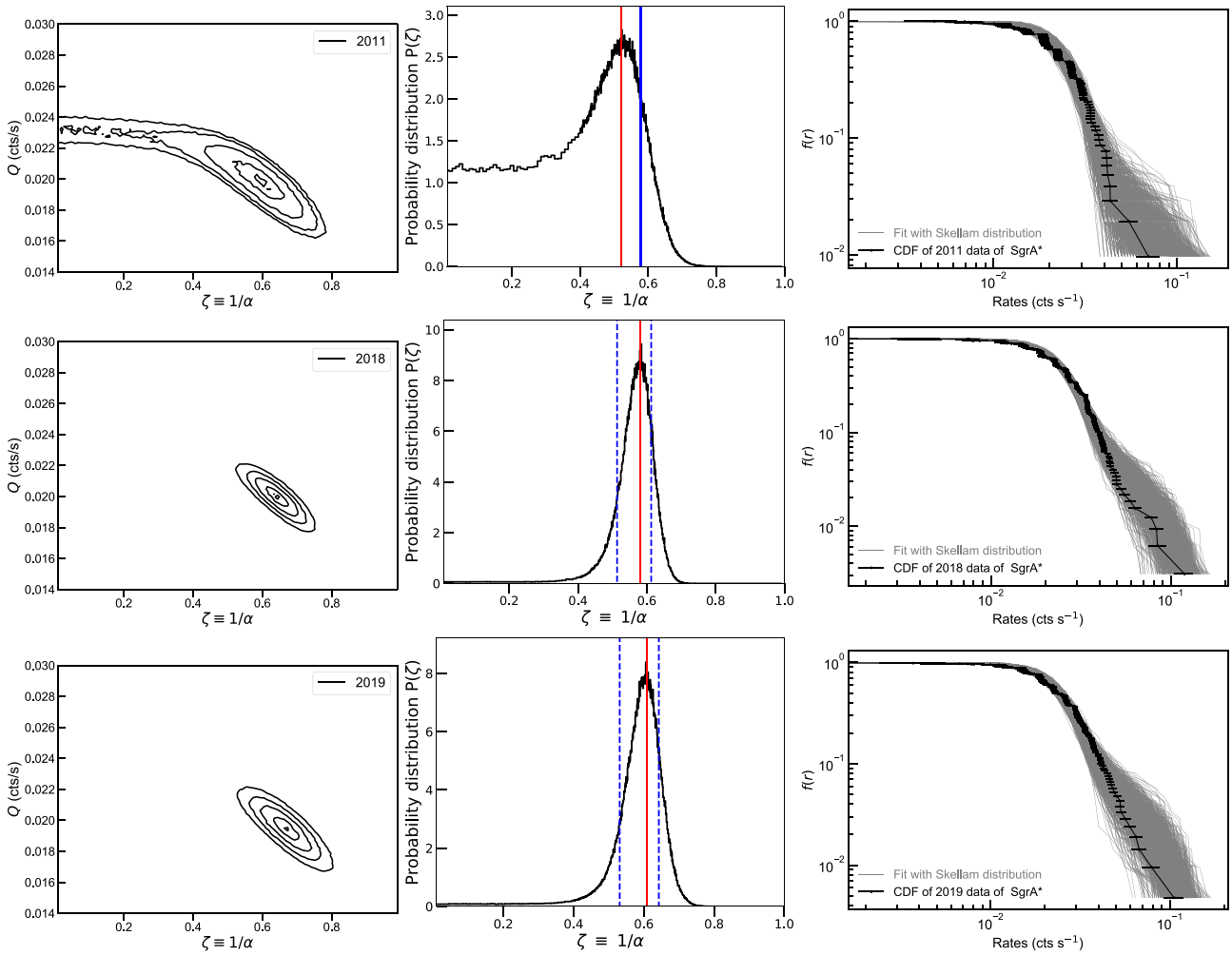

Figure B1 *– continued*

This paper has been typeset from a TEX/LATEX file prepared by the author.